

CLAIMS:

Having thus described our invention, what we claim as new, and desire to secure by Letters Patent is:

1. A one-bounce data network comprising a plurality of nodes interconnected to each other via communication links, the network including a plurality of interconnected switch devices, said switch devices interconnected such that a message is communicated between any two switches passes over a single link from a source switch to a destination switch; and, the source switch concurrently sends a message to an arbitrary bounce switch which then sends the message to the destination switch.
2. The one-bounce data network as claimed in Claim 1, wherein all the switches in the network are paired and simultaneously communicate, such that each pair communicates at half the aggregate bandwidth of all the links from a switch, whereby at each switch of a pair, one-half of the bandwidth serves the bounce message of the other switch of said pair.
3. The one-bounce data network as claimed in Claim 2, wherein the effective bandwidth from each switch increases from one-half aggregate up to all the aggregate bandwidth of all the links from the switch as a number of switches comprising a group increases in number, the largest group comprising all the switches of the network.
4. The one-bounce data network as claimed in Claim 1, wherein the network is circuit-switched.
5. The one-bounce data network as claimed in Claim 1, wherein the network is packet-based, wherein a packet to be communicated includes means for indicating whether a packet has already bounced or not.

6. The one-bounce data network as claimed in Claim 5, wherein the indicating means comprises a bounce bit in a packet header that is reset when the packet is injected into the network and is set when the packet is bounced to said arbitrary other switch, a set bounce bit indicating that the packet can only go to the destination switch.

7. The one-bounce data network as claimed in Claim 6, further including means for enabling a node to specify a particular bounce switch to be used for message communication, a node setting the bounce bit in the packet header and specifying a direct router identifier, wherein when the identified direct router is not the router of a destination node, the packet is bounced over an external link to the identified direct router.

8. The one-bounce data network as claimed in Claim 1, wherein each switch includes three independent channels for communicating message traffic internal to said switch free of deadlock, said independent channels being used in concert to provide a single deadlock-free channel across the switches rendering said network as deadlock free.

9. The one-bounce data network as claimed in Claim 6, wherein said bounce bit is set before injecting a packet into the network, such that said packet is guaranteed not to bounce, said packet injected for communication on an injection channel in an injection switch and delivered to the destination channel on the destination switch, a delivery being in order across the network.

10. The one-bounce data network as claimed in Claim 1, wherein a switch comprises one or more routers, internal links and external links, an internal link linking two routers within the switch and the external link linking two routers in two different switches wherein, for a switch, a router may have zero, one or more external links.

11. The one-bounce data network as claimed in Claim 10, wherein within a given switch, the internal links may have differing bandwidths.

12. The one-bounce data network as claimed in Claim 10, wherein within a given switch, the external links have different bandwidths.
13. The one-bounce data network as claimed in Claim 10, wherein within a given switch, the external links have different bandwidths than those of the internal links, such different bandwidths matching the aggregate performance of the internal and external links.
14. The one-bounce data network as claimed in Claim 10, wherein a switch is integrated such that an internal router serves the external links in addition also serving the nodes of the switch.
15. The one-bounce data network as claimed in Claim 10, wherein a switch is a two-part switch having a first group of internal routers serving the external links and a second group of routers serving the nodes of the switch.
16. The one-bounce data network as claimed in Claim 10, wherein said network is configured to perform all-to-all messaging, wherein each router has an external link, and is configured with internal links having twice the bandwidth of the external link.
17. The one-bounce data network as claimed in Claim 10, wherein switches of said network is internally a multi-dimensional torus network.
18. The one-bounce data network as claimed in Claim 10, wherein switches of said network is internally a multi-dimensional switch network.
19. The one-bounce data network as claimed in Claim 10, configured as one of: a single-level, two-level or multi-level bounce network.
20. The one-bounce data network as claimed in Claim 10, further comprising means for injecting a unicast packet on the network and through a injection switch, said unicast packet including a first field in a packet header for specifying a destination node identifier and another

field in said packet header for specifying a direct router, wherein a node writes said destination node identifier into the packet header before injecting a packet into the network at said injection switch, and said router implementing means for determining a direct router on the injection switch with an external link to the destination switch based on said destination node identifier specified.

21. The one-bounce data network as claimed in Claim 20, wherein said direct router identifier is used within the injection switch to route a packet along internal links to said direct router having an external link to a destination switch, a packet possibly encountering other routers en route to said direct router, wherein an adaptive routing means at a router determines if a packet being routed should be bounced to another switch at each router encountered having one or more external links to other switches, and upon determining a packet is to be bounced, setting said bounce bit.

22. A distributed-memory computer configured as a two-level one-bounce network comprising: a plurality of cards, each card including 64 routers connected as a 4*4*4 torus network, with each router connected to a single node.

23. The distributed-memory computer as claimed in Claim 22, wherein a rack unit comprises 16 cards, the 16 cards in a rack being connected as a rack-level one-bounce network with each card connected by 4 rack-level links to each other card in its rack.

24. The distributed-memory computer as claimed in Claim 23, comprising 64 racks, the 64 racks in the machine being connected as a machine-level one-bounce network with each rack connected by 16 machine-level links to each other rack.

25. A distributed-memory computer configured as a binary one-bounce network comprising L levels, said network comprising a total number of 2^x routers wherein $x = 2^{(L-1)}$, a L-level of said network having L links at each router, wherein a given level includes double the link bandwidth compared to bandwidth at a next lower level.

26. A method of passing messages over a one-bounce network comprising a plurality of nodes interconnected to each other via communication links, the network including a plurality of interconnected switch devices, said switch devices interconnected such that a message communicated between any two switches passes over a single link from a source switch to a destination switch said method comprising the steps of:

concurrently communicating a message from said source switch to an arbitrary other switch, and

communicating the message from the arbitrary other switch to the destination switch.

27. The method as claimed in Claim 26, including the step of configuring switches in said network as pairs, such that each pair communicates at half the aggregate bandwidth of all the links from a switch, whereby at each switch of a pair, one-half of the bandwidth serves the bounce message of the other switch of said pair.

28. The method as claimed in Claim 27, further including the step of increasing the number of switches comprising a group such that the effective bandwidth from each switch increases from one-half aggregate up to all the aggregate bandwidth of all the links from the switch, a largest group comprising all the switches of the network.

29. The method as claimed in Claim 26, wherein the network is packet-based, said method including injecting a packet into a network switch and indicating whether a packet has bounced or not.

30. The method as claimed in Claim 29, wherein the step of indicating comprises resetting a bounce bit in a packet header when the packet is injected into the network and setting the bounce bit when the packet is bounced to said arbitrary other switch, a set bounce bit indicating that the packet can only go to the destination switch.

31. The method as claimed in Claim 29, further including the step of: enabling a node to specify a particular bounce switch to be used for message communication, a node setting the bounce bit in the packet header and specifying a direct router identifier, wherein when the identified direct router is not the router of a destination node, the packet is bounced over an external link to the identified direct router.
32. The method as claimed in Claim 30, wherein each switch includes three independent channels for communicating message traffic internal to said switch free of deadlock, said method including implementing said channels to provide a single deadlock-free channel across the switches rendering said network as deadlock free.
33. The method as claimed in Claim 32, further including the step of: setting said bounce bit before injecting a packet into the network, such that said packet is guaranteed not to bounce, said packet injected for communication on an injection channel in an injection switch and delivered to the destination channel on the destination switch, a delivery being in order across the network.
34. The method as claimed in Claim 26, wherein a switch comprises one or more routers, internal links and external links, an internal link linking two routers within the switch and the external link linking two routers in two different switches wherein, for a switch, a router may have zero, one or more external links.
35. The method as claimed in Claim 34, wherein within a given switch, one of: the internal links have differing bandwidths, the external links have different bandwidths or, the external links have different bandwidths than those of the internal links, such different bandwidths matching the aggregate performance of the internal and external links.
36. The method as claimed in Claim 35, wherein a switch is integrated such that an internal router serves the external links in addition also serving the nodes of the switch.

37. The method as claimed in Claim 35, wherein a switch is a two-part switch having a first group of internal routers serving the external links and a second group of routers serving the nodes of the switch.
38. The method as claimed in Claim 35, comprising the step of: configuring said network to perform all-to-all messaging, wherein each router has an external link, and is configured with internal links having twice the bandwidth of the external link.
39. The method as claimed in Claim 37, wherein internal switches of said network is one of a multi-dimensional torus network or, a multi-dimensional switch network.
40. The method as claimed in Claim 34, further comprising the steps of: injecting a unicast packet on the network and through a injection switch, said unicast packet including a first field in a packet header for specifying a destination node identifier and another field in said packet header for specifying a direct router, wherein a node writes said destination node identifier into the packet header before injecting a packet into the network at said injection switch, and said router implementing means for determining a direct router on the injection switch with an external link to the destination switch based on said destination node specified.
41. The method as claimed in Claim 40, further including routing a packet along internal links to said direct router having an external link to a destination switch according to said direct router identifier, a packet possibly encountering other routers en route to said direct router, said method further including determining at an encountered router if a packet being routed should be bounced to another switch having one or more external links to other switches, and upon determining a packet is to be bounced, setting said bounce bit.